Molecular & Biochemical Parasitology 172 (2010) 31–40

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

Transcript analysis of parasitic females of the sedentary semi-endoparasitic nematode *Rotylenchulus reniformis*☆

Martin J. Wubben[a,*], Franklin E. Callahan[a], Brian S. Scheffler[b]

[a] USDA/ARS, Crop Science Research Laboratory, Mississippi State, MS 39762, USA
[b] USDA/ARS, Mid-South Area Genomics Laboratory, Stoneville, MS 38776, USA

ARTICLE INFO

Article history:
Received 10 February 2010
Received in revised form 12 March 2010
Accepted 15 March 2010
Available online 24 March 2010

Keywords:
Reniform nematode
EST
Sequencing
Parasitism
RNAi
Plant

ABSTRACT

Rotylenchulus reniformis, the reniform nematode, is a sedentary semi-endoparasitic nematode capable of infecting >300 plant species, including a large number of crops such as cotton, soybean, and pineapple. In contrast to other economically important plant-parasitic nematodes, molecular genetic data regarding the *R. reniformis* transcriptome is virtually nonexistant. Herein, we present a survey of *R. reniformis* ESTs that were sequenced from a sedentary parasitic female cDNA library. Cluster analysis of 2004 high quality ESTs produced 123 contigs and 508 singletons for a total of 631 *R. reniformis* unigenes. BLASTX analyses revealed that 39% of all unigenes showed similarity to known proteins ($E \leq 1.0e{-}04$). *R. reniformis* genes homologous to known parasitism genes were identified and included β-1,4-endoglucanase, fatty acid- and retinol-binding proteins, and an esophageal gland cell-specific gene from *Heterodera glycines*. Furthermore, a putative ortholog of an enzyme involved in thiamin biosynthesis, thought to exist solely in prokaryotes, fungi, and plants, was identified. Lastly, 114 *R. reniformis* unigenes orthologous to RNAi-lethal *Caenorhabditis elegans* genes were discovered. The work described here offers a glimpse into the transcriptome of a sedentary semi-endoparasitic nematode which (i) provides the transcript sequence data necessary for investigating engineered resistance against *R. reniformis* and (ii) hints at the existance of a thiamin biosynthesis pathway in an animal.

Published by Elsevier B.V.

1. Introduction

The reniform nematode, *Rotylenchulus reniformis* Linford & Oliveira, is an obligate, sedentary, semi-endoparasitic root pathogen of more than 300 plant species [1]. Included within the *R. reniformis* host range are a number of economically important crops such as cotton, soybean, pineapple, tomato, and sweet potato [1,2]. The economic impact exerted by *R. reniformis* on Upland cotton production is particularly severe with annual losses of ca. $230 million in the United States [3–5]. While natural resistance to *R. reniformis* has been incorporated into elite soybean germplasm [6], there are currently no Upland cotton cultivars resistant to this nematode [4,5]. The *R. reniformis* life-cycle begins when second-stage juveniles (J_2) hatch from eggs in the soil or from eggs within the gelatinous matrix associated with adult sedentary females. In contrast to sedentary endoparasites like cyst (*Heterodera* and *Globodera* spp.) and root-knot nematode (*Meloidogyne* spp.), *R. reniformis* J_2 do not immediately infect the host root, but instead become inactive and progress through three molting events in the absence of feeding [1]. After the final molt, vermiform, infective adult females seek out a host and penetrate the root epidermis; adult males do not feed. The posterior half of the female remains outside the root and eventually swells to give the characteristic kidney shape. The feeding site is established on endodermal and pericycle cells adjacent to the vasculature by the vermiform female and closely resembles syncytia formed by cyst nematodes [1,2]. After fertilization, 30–200 eggs are deposited into a gelatinous matrix produced by the mature reniform-shaped female [1,2].

The signaling events leading to feeding site formation by sedentary plant-parasitic nematodes are believed to be mediated by the products of nematode parasitism genes which are secreted by the esophageal gland cells and injected into host root cells via the nematode stylet [7–9]. The first parasitism genes identified encoded a battery of cell wall degrading or loosening proteins such as β-1,4-endoglucanase, pectate lyase, expansin, and cellulose binding proteins [9]. In addition to plant cell wall manipulation, parasitism genes have been discovered that encode chorismate mutase [10], SKP1-like proteins [11], and plant CLE-like proteins [12]; however, the majority of parasitism genes, regardless of species, lack similarity to known proteins and perform unknown functions [7,8]. Despite being primarily 'pioneer' sequences, para-

☆ Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the United States Department of Agriculture.

* Corresponding author. Tel.: +1 662 320 7585.
E-mail address: martin.wubben@ars.usda.gov (M.J. Wubben).

0166-6851/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.molbiopara.2010.03.011

sitism gene candidates have been identified from large EST datasets based on the presence of a signal peptide in the predicted protein coupled with the lack of transmembrane domains [13].

The importance of individual parasitism genes to feeding site development has been studied using RNA-interference (RNAi). RNAi is an ubiquitous phenomenon in eukaryotes that involves the regulation of gene expression within a cell or tissue by RNA molecules complementary to specific mRNA species [14]. The targeted silencing of parasitism genes using RNAi has been shown to significantly reduce nematode infection and/or reproduction on a susceptible host [15–18]. In addition to parasitism genes, RNAi-mediated silencing of highly conserved genes with roles in vital cellular or developmental processes has also shown promise in reducing nematode infection [15,19–22].

The potential shown by RNAi-mediated gene silencing for engineering nematode resistance in crops rests on a foundation of functional genomic data for the parasite(s) in question. For example, almost 80,000 ESTs from *Meloidogyne* spp. are publicly available [23]. More than 43,000 ESTs from cyst nematodes and thousands more from lesser known species such as *Radopholus similis* [24], *Ditylenchus africanus* [25], *Bursaphelenchus* spp. [26,27] and *Pratylenchus* spp. [28] have been made available. In stark contrast, as of August 2009, only 56 *R. reniformis* ESTs were listed in Genbank; however, 55/56 were ribosomal sequences, leaving a single EST that corresponded to a cathepsin L-like cysteine proteinase (gb|AY999066.1|). It is obvious that this severe lack of cDNA sequence information must be addressed before any RNAi-related research toward engineered *R. reniformis* resistance can be pursued.

Herein, we describe the construction of a cDNA library from sedentary parasitic females of *R. reniformis* and the analysis of 2004 ESTs from that library. Clustering yielded a total of 631 *R. reniformis* unigene sequences. A thorough analysis of these unigenes revealed a number of potential parasitism genes, including a putative cellulase, as well as genes highly conserved in *C. elegans* that show severe RNAi-mediated gene silencing phenotypes. Furthermore, an *R. reniformis* unigene was identified that may represent an enzyme involved in vitamin B1 biosynthesis which is thought to occur exclusively in fungi, plants, and prokaryotes.

2. Materials and methods

2.1. Isolation of R. reniformis sedentary females

An inbred *R. reniformis* population was maintained on cotton plants in sand in a greenhouse. This culture had originally been acquired from Dr. Forrest Robinson, USDA/ARS (retired), College Station, TX. Infected cotton roots from 3 to 5 plants were washed free of sand with tap-water. Lateral roots were removed from the taproot with scissors and placed in a Waring blender which was then filled with just enough water to cover all the roots. The roots were macerated by two 5-s pulses on the "LOW" setting. Detached sedentary females were collected by sieving the root slurry through a #60/#100 sieve stack (U.S. Standard Sieve Series, Fisher Scientific, Pittsburgh, PA, USA). The #100 sieve material was re-sieved through a #80/#100 sieve stack to remove additional root material. Then, the #100 sieve material was transferred to a 250 mL beaker. Female nematodes were allowed to settle to the bottom of the beaker, and the supernatant, containing fine root material, was decanted and discarded. Root and sand material that persisted was removed by centrifugal sucrose-floation [29]. Female nematodes were collected from the water/sucrose interphase with a glass Pasteur pipet, transferred to a 50 mL tube (Greiner Bio-One, Monroe, NC, USA) and washed with sterile Milli-Q water (Millipore, Bedford, MA, USA). Any remaining debris was removed by pouring the nematodes into a petri dish and picking out the debris with a Pasteur pipet. Nematode viability was verified at this time by observing pharyngeal pumping using a Nikon SMZ1500 stereomicroscope (Nikon Inc., Melville, NY, USA). Nematodes were then washed three times with sterile Milli-Q water before being transferred to a 1.5 mL centrifuge tube, pelleted, and flash frozen in liquid nitrogen.

2.2. Total RNA isolation and cDNA library construction

Female nematodes were ground in liquid nitrogen using a blue plastic pestle fitted to a 1.5 mL centrifuge tube. Total RNA was isolated using the animal tissue protocol that accompanied the RNeasy Plus Mini Kit (Qiagen Inc., Valencia, CA, USA). Contaminating genomic DNA was removed using the DNA-*free* Kit according to the manufacturer's instructions (Ambion, Austin, TX, USA). A cDNA library was constructed from 150 ng of DNase-treated total RNA using the Creator SMART cDNA Library Kit (Clontech, Mountain View, CA, USA). cDNA amplification, size fractionation, restriction enzyme digestion, ligation into the pDNR-LIB vector, and *Escherichia coli* transformation was performed as outlined in the kit manual including the use of all positive and negative controls. Library titre was estimated to be 1.64×10^5 cfu/mL. Colony PCR of 13 randomly selected clones gave an average insert size of ≈500 nt.

2.3. Clone sequencing and bioinformatic analyses

Recombinant clones were identified by blue/white selection on Luria Bertani (LB) agar plates containing X-gal, IPTG, and 30 μg/mL chloramphenicol. Blue colonies were picked into 96-well V-bottom microplates (Fisher Scientific) containing 150 μL LB broth + chloramphicol. Plates were sealed and incubated overnight with shaking at 37 °C. Sterile glycerol was added to each well to reach a final concentration of 15%. 75 μL from each well was transferred to a new 96-well plate which was sealed with tape and shipped on dry ice to the USDA/ARS MidSouth Area (MSA) Genomics Facility. Sequencing reactions were performed using the BigDye Terminator Cycle Sequencing Reaction Kit version 3.1 (Applied Biosystems) and detected using the Applied Biosystems 3730XL DNA Analyzer.

Vector and low quality sequences were trimmed from the raw sequence reads using Phred [30,31] and Lucy [32]. Sequence assembly was performed using Sequencher v4.7™ software (Gene Codes Corporation, Ann Arbor, MI, USA) under the "dirty data" option with 90% minimum shared identity and a minimum sequence overlap of 40 nt. Unigene sequence database searches were performed using the Basic Local Alignment Search Tool (BLAST) under default parameters [33]. Unigene sequence annotations were performed by combining Gene Ontology mapping and InterProScan data using Blast2Go software [34].

3. Results

3.1. Library characteristics and cluster analysis

2784 clones from an *R. reniformis* parasitic female cDNA library were picked and bi-directionally sequenced using standard M13 forward and reverse primers. General library characteristics and results of EST cluster analysis are presented in Table 1. Clones that yielded ESTs of poor quality ($n = 478$), ribosomal origin ($n = 272$), or mitochondrial origin ($n = 30$) were dropped from the analysis, leaving 2004 ESTs for further study. Prior to clustering, the mean length of all ESTs was 423.7 nucleotides (nt). To reduce dataset redundancy, ESTs were grouped into contigs using Sequencher™ v4.7 based on shared sequence identity (≥90%) and a minimum overlap of 40 nt. Cluster analysis yielded 123 contigs (mean length = 539.1 nt) and 508 singletons (mean length = 425.5 nt) for a

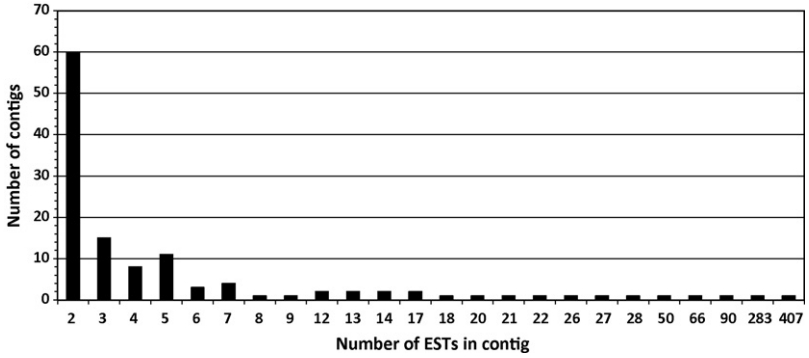

Fig. 1. Distribution of *Rotylenchulus reniformis* contig sizes resulting from cluster analysis of 2004 expressed sequence tags from sedentary parasitic females based on ≥90% nucleotide identity and a 40 nt minimum overlap.

total of 631 unigenes. The number of ESTs within a contig ranged from 2 to 407; however, the vast majority of contigs fell within the 2–5 EST range (Fig. 1). Given a conservative estimate of 17,300 genes for the model species *C. elegans* [35] and 19,212 genes for the plant-parasite *M. incognita* [36], the 631 *R. reniformis* unigenes described here likely account for ≈3.5% of the reniform nematode transcriptome.

3.2. BLAST analyses

All *R. reniformis* unigenes were compared to the NCBI non-redundant database (NRdb) using BLASTN and BLASTX where a significant match between the query and subject sequences was defined as having an expected value (E) ≤ 1.0e−04. BLASTN identified significant subject matches for 143/631 unigenes (22.7%) with a mean top hit E = 8.3e−06. Decreasing the E-value threshold to 1.0e−10 decreased the number of unigenes with significant subject matches to 93/631 (14.7%). A large portion of the most significant unigene matches (E ≤ 1.3e−73) were to orthologs of conserved eukaryotic genes, e.g., histones H2 and H4, RNA polymerase II, translation initiation and elongation factors, and cytoskeletal proteins. Highly significant matches to plant-parasitic nematode genes were also identified; for example, the unigenes Contig106, FRC-1C_M13F_G04, and Contig111 matched *H. glycines* orthologs for a class V aminotransferase (E = 1.8e−123), an arginine kinase (E = 1.3e−116), and an aldolase (E = 1.6e−64), respectively. Unigene FRC-4D_M13F_A12 showed significant similarity to the *G. pallida* SEC2 gene (5.5e−61), which is believed to be involved in potato cyst nematode parasitism [37,38].

BLASTX identified significant subject matches for 246/631 (39.0%) unigenes; however, this value decreased to 168 unigenes at an E-value threshold of 1.0e−10. Taking these figures into account, 385–463 unigenes (61–73%) potentially represent novel proteins. Proteins from the human nematode parasite, *Brugia malayi*, accounted for 30.0% of the top BLASTX matches. Proteins from the model free-living nematode species *C. elegans* and *C. briggsae* accounted for an additional 27.2% of the top BLASTX matches. Only 16 unigenes (6.5%) were most similar to plant-parasitic nematode proteins with all 16 corresponding to *H. glycines*. Six unigenes showed top BLASTX matches to prokaryotic proteins; however, for five of these six unigenes, significant matches to eukaryotic proteins were also identified. Unigene FRC-3A_D06 showed matches to only prokaryotic isochorismatase hydrolase proteins. Contig218 showed a high level of similarity to an ortholog of NMT1 from the fungus *Gibberella zeae* (E = 7.0e−86). NMT1, also known as THI5 in the yeast *Sacchromyces cereviseae*, is required for thiamin (vitamin B1) biosynthesis in fungi and plants [39]; however, NMT1-like genes have not been identified in other eukaryotic organisms.

A closer examination of the most significant BLASTX matches revealed that there were groups of unigenes that shared identical subject matches (Table 2). Nineteen groups were identified with each group consisting of 2–4 unigenes. One explanation for this finding is that unigenes within a group represent transcripts from a single locus; however, the unigenes share <90% nucleotide identity and they align to identical regions of the subject sequence which does not suggest an instance of under-clustering or fragmentation [40]. Because *R. reniformis* is amphimictic and the genetic variation within the population used for library construction was unknown, it is possible that unigenes within a group represent different alleles of a locus or multiple loci or are products of alternative splicing events. Evidence of alternative splicing was found within Groups 16 and 17 (Table 2). In Group 16, we determined that Contig10 and Contig416 shared almost 100% identity; however, Contig416 showed a 249 nt in-frame sequence gap relative to Contig10. Likewise, Contig11 and FRC-7B_M13F_C08 shared 98% sequence identity but with an 18 nt in-frame gap (Table 2). Similar relationships between unigenes within the remaining 17 groups were not detected.

Our cDNA library represented a non-normalized collection of *R. reniformis* transcript fragments; therefore, inferences about the relative expression of different genes could be made based on the relative abundance of their corresponding ESTs. Following this rationale, the 10 most abundant unigene sequences are presented in Table 3. Contig136 was most prevalent, consisting of 407 ESTs. Unfortunately, no identifiable homologs for Contig136 could be found by any BLAST program; however, the sequence showed 84.1% A/T composition which may suggest a mitochondrial origin for this unigene [24]. Various serine-type endopeptidase inhibitor proteins represented the second, fourth, fifth, seventh, and ninth most abundant sequences in the library. Collectively, this sequence type was comprised of 490 ESTs or 24.5% of all ESTs. The third most abundant sequence, as represented by Group 16 (143 ESTs), was most similar to an *H. glycines* C-type lectin protein.

Table 1
Rotylenchulus reniformis cDNA library characteristics.

Clones sequenced	2784
Poor quality sequences (e.g. <100 nt)	478 (17.2%)
Ribosomal sequences	272 (9.8%)
Mitochondrial sequences	30 (1.1%)
Sequences eligible for cluster analysis	2004
Mean sequence length (nt)	423.7
Number of contigs	123
Mean contig length (nt)	539.1
Number of singletons	508
Mean singleton length (nt)	425.5
Total number of consensus sequences	631

nt: nucleotides.

Table 2
Grouping of *Rotylenchulus reniformis* consensus sequences by putative function according to shared top BLASTx hits.

Group	Consensus sequence	GenBank accession no.	No. ESTs[a]	BLASTx hit description	*E*-Value	Hit accession no.	Organism
1	Contig1_76_406	EZ421800	20	Serine peptidase inhibitor-with	4.2e−06	P84875	*Sabellastarte magnifica*
	Contig224	EZ421847	12	kunitz and wap domains 1	7.2e−06		
	Contig395	EZ421884	28		5.5e−06		
2	Contig106	EZ421807	2	Alanine-glyoxylate	1.1e−88	AAK26375	*Heterodera glycines*
	FRC-01-test_C11	GT736460	1	aminotransferase	7.6e−32		
3	FRC-1B_M13F_A04	GT736596	1	Testis expressed gene 2	5.1e−34	XP_001895126	*Brugia malayi*
	FRC-5A_M13F_C11	GT737639	1		3.4e−16		
4	FRC-1A_M13F_D05	GT736550	1	Cysteine-rich venom protein	3.4e−08	XP_001674547	*Caenorhabditis briggsae*
	FRC-7B_A09	GT738255	1		5.0e−07		
5	Contig3	EZ421860	13	Serine peptidase inhibitor-with	2.0e−08	NP_001102927	*Rattus norvegicus*
	Contig301	EZ421862	2	kunitz and wap domains 1	7.6e−08		
	Contig5	EZ421903	2		3.8e−07		
6	Contig8_211	EZ421917	4	h4 histone protein	7.3e−39	XP_001898784	*Brugia malayi*
	FRC-6A_G05	GT737930	1		7.2e−39		
	FRC-7C_D07	GT738328	1		7.2e−38		
7	Contig101	EZ421803	2	srf-type transcription factor	4.5e−12	ZP_02925926	*Verrucomicrobium spinosum*
	FRC-1C_M13F_C10	GT736668	1		4.2e−14		
8	Contig479	EZ421901	2	Chondroitin ProteoGlycan family	1.7e−12	NP_492047	*Caenorhabditis elegans*
	Contig87	EZ421920	2	member (cpg-3)	1.6e−12		
9	Contig411	EZ421889	2	Vitellogenin precursor-like protein	2.1e−05	CAK18212	*Ascaris suum*
	Contig44_54	EZ421895	2		2.7e−04		
	FRC-7C_A08	GT738324	1		1.0e−05		
10	Contig126	EZ421818	4	Kunitz inhibitor-like protein 2	2.5e−06	XP_001983642	*Drosophila grimshawi*
	FRC-5B_A07	GT737704	1		2.4e−06		
11	Contig232_234	EZ421850	5	Kunitz inhibitor-like protein 2	2.4e−06	XP_002070501	*Drosophila willistoni*
	Contig322	EZ421869	2		2.7e−05		
	FRC-3D_G03	GT737272	1		2.9e−07		
	Contig170	EZ421830	8		7.5e−06		
12	Contig226	EZ421848	283	Similar to papilin	3.2e−06	XP_002121111	*Ciona intestinalis*
	Contig27	EZ421854	14		3.2e−06		
	Contig370	EZ421877	6		7.2e−06		
13	Contig15	EZ421824	66	Tissue factor pathway inhibitor	4.1e−07	NP_001128730	*Sus scrofa*
	Contig20	EZ421836	22		6.6e−07		
14	Contig78	EZ421915	2	Hypothetical protein	7.3e−04	ACI49104	*Caenorhabditis brenneri*
	FRC-7A_M13F_H02	GT738227	1	Cbre_JD13.006	8.2e−04		
15	Contig97	EZ421921	2	Pyruvate kinase	8.5e−28	CAV31772	*Caenorhabditis elegans*
	FRC-01-test_D07	GT736461	1		3.6e−63		
16	Contig10[b]	EZ421801	90	c-type lectin	7.4e−15	AAK94488	*Heterodera glycines*
	Contig357	EZ421875	27		7.9e−15		
	Contig416[b]	EZ421891	26		5.0e−15		
17	Contig11[b]	EZ421808	21	Putative secreted protease	2.1e−12	AAM93648	*Ixodes scapularis*
	FRC-7B_M13F_C08[b]	GT738279	1	inhibitor	2.1e−13		
18	Contig150	EZ421825	5	Protein partially confirmed by	2.7e−08	CAX65088	*Caenorhabditis elegans*
	FRC-3D_A01	GT737268	1	transcript evidence	7.8e−08		
	FRC-6D_M13F_D03	GT738139	1		6.0e−05		
19	Contig116	EZ421812	3	Putative gland protein G22C12	6.6e−10	AAP30768	*Heterodera glycines*
	Contig213	EZ421843	2		3.2e−07		
	FRC-3C_M13R_E06	GT737254	1		2.1e−05		
	FRC-5D_B11	GT737857	1		1.2e−07		

[a] Number of expressed sequence tags (ESTs) that correspond to the consensus sequence.
[b] Within a group, these consensus sequences may represent different splice isoforms transcribed from a single locus

3.3. Annotation of *R. reniformis* unigenes

The *R. reniformis* unigene sequences were annotated using the Blast2GO software package [34] according to the universal GO terminology for cellular compartment, biological process, and molecular function. We found that 231 unigenes had at least one GO term associated with them. The number of GO terms associated with a particular unigene ranged from 1 to 31 with 119 unigenes having 1–3 GO terms. Regarding the cellular compartment categorization, 17% of unigenes were predicted to be part of a protein complex (Fig. 2). Equivalent numbers of unigenes were predicted to be targeted to the nucleus or to the extracellular region (Fig. 2). The high number of extracellular proteins is likely due to the many predicted protease inhibitor proteins in our dataset. This fact would also explain why serine-type endopeptidase inhibitor activity was by far the most prevalent molecular function identified; encompassing 42% of the annotated sequences (Fig. 2). In contrast, no single biological process category appeared significantly larger

Table 3
The 10 most abundant sequences identified from the *Rotylenchulus reniformis* sedentary female cDNA library.

Consensus sequence(s)[a]	GenBank accession no.(s)	# of ESTs (% of total)[b]	Top BLASTx hit description	SignalP[c]	TMHMM[d]
Contig136	EZ421822	407 (20.3)	No hits	Negative	Negative
Group 12	EZ421848 EZ421854 EZ421877	303 (15.1)	gi|198418054|ref|XP_002121111.1| PREDICTED: similar to papilin [*Ciona intestinalis*] (3.2e−06)	Positive	Positive
Group 16	EZ421801 EZ421875 EZ421891	143 (7.1)	gi|22073913|gb|AAK94488.1| putative C-type lectin [*Heterodera glycines*] (7.4e−15)	Positive	Negative
Group 13	EZ421824 EZ421836	88 (4.4)	gi|158668327|gb|ABW76681.1| tissue factor pathway inhibitor [*Sus scrofa*] (4.1e−07)	Positive	Positive
Group 1	EZ421800 EZ421847 EZ421884	60 (3.0)	gi|118196994|emb|CAK55547.1| putative carboxypeptidase inhibitor [*Sabellastarte magnifica*] (5.5e−06)	Positive	Negative
Contig4	EZ421886	50 (2.5)	gi|71992017|ref|NP_499406.2| hypothetical protein W05B2.2 [*Caenorhabditis elegans*] (1.0e−04)	Positive	Positive
Group 17	EZ421808 GT738279	22 (1.1)	gi|22164286| putative secreted protease inhibitor ISL1156 cluster 318 [*Ixodes scapularis*] (2.1e−12)	Positive	Negative
Contig105_195	EZ421806	18 (0.9)	No hits	Positive	Positive
Contig272	EZ421855	17 (0.8)	gi|110756487| PREDICTED: similar to Papilin CG33103-PB isoform B isoform 1 [*Apis mellifera*] (8.6e−07)	Positive	Positive
Contig21	EZ421842	17 (0.8)	No hits	Positive	Positive

[a] Either individual contigs comprised of overlapping ESTs or Groups of contigs and/or singletons that share the same top BLASTx hits as shown in Table 2.
[b] Total number of overlapping ESTs that comprise the individual contig or the sum total of ESTs that comprise the respective members of a Group.
[c] Predicted peptide(s) tested positive or negative for a signal peptide according to SignalP 3.0 [55].
[d] Predicted peptide(s) tested positive or negative for the presence of transmembrane helices according to TMHMM 2.0 [56].

than any other; however, three categories describing metabolic and biosynthetic processes accounted for a combined 36% of annotated unigenes (Fig. 2).

3.4. TBLASTX analysis of R. reniformis unigenes

In order to identify as many potential homologs as possible, *R. reniformis* consensus sequences were compared to the NCBI est_others database using TBLASTX. This analysis showed that 246 sequences (38.9%) identified with at least one other nematode EST in Genbank (E_{max} = 9.40e−05). The genera of nematodes corresponding to the top TBLASTX matches were recorded and the total number of top hits for each genus determined (Fig. 3). Various cyst nematode species were the most common plant parasites encountered with *H. glycines* and *G. pallida* being most prevalent (Fig. 3). Root-knot forming nematode species (*Meloidogyne* spp.) collectively formed the next most common group (Fig. 3). These observations may not be surprising because both cyst and root-knot nematodes are sedentary parasites like *R. reniformis*. The number of top TBLASTX matches to proteins from migratory plant-parasitic nematodes, e.g., *R. similis* and *Pratylenchus* spp., was slightly larger than the number of *Caenorhabditis* spp. top hits (Fig. 3). Top TBLASTX hits to animal-parasitic nematode species were also observed, albeit rarely, and included *Ancylostoma caninum*, *Ascaris suum*, *Brugia* spp., and *Toxocara canis*.

3.5. Identification of candidate parasitism genes

Our BLASTX analysis of the *R. reniformis* unigenes identified seven cDNAs whose only matches were to proteins from other plant- or animal-parasitic nematodes (Table 4). Sequences Contig116, Contig46, and FRC-5A_M13F_D07 showed homology to the previously identified parasitism genes gland protein G22C12 (*H. glycines*), SEC-2 (*G. pallida*), and cellulase (Table 4). A defining characteristic of parasitism genes is the presence of a signal peptide in the absence of helical transmembrane domains [8,13,41]. The predicted proteins of Contig116, FRC-1B_E06, and FRC-2C_M13R_G08 showed these characteristics (Table 4); however, it should be noted that our unigene sequences may not contain full-length open reading frames (ORF) or they may lack 5′ ends, leading to false negative results for the presence of a transmembrane domain or signal peptide. An example of this scenario would be FRC-5A_M13F_D07, a homolog of other nematode cellulase genes which are well-known to be secreted proteins [8]. Indeed, upon closer examination, it was clear that FRC-5A_M13F_D07 corresponded to the 3′-end of the putative cellulase gene and lacked the N-terminus of the predicted protein.

Thus far, the majority of parasitism genes that have been identified from plant-parasitic nematodes lack homology with known proteins [13]. Within our dataset, 29 *R. reniformis* cDNAs were identified that showed no BLASTN or BLASTX homology with known sequences but were predicted to encode proteins having a signal peptide but no transmembrane domain (Table 5). Twenty-six sequences appeared to have complete ORFs that were predicted to encode secreted proteins ranging in length from 40 to 296 amino acids (aa); however, 21 of these 26 predicted proteins were ≤120 aa (Table 5). TBLASTX detected weak matches between six sequences and ESTs of other plant-parasitic nematodes; whereas, Contig413 aligned with an EST from artichoke (Table 5). In addition to TBLASTX, predicted peptide sequences were analyzed via BLASTP. Only three of 29 predicted proteins showed significant homology to known proteins (Table 5). This included Contig320, which showed similarity to the *H. glycines* putative gland protein G16B09 (Table 5). Oddly, signal peptides were detected from two different ORFs for FRC-4D_G02 and FRC-7C_M13F_G05.

3.6. Homology to Caenorhabditis elegans

Each *R. reniformis* unigene sequence was compared against the *C. elegans* WS201 protein list using BLASTX. We found that 247 sequences showed significant similarity ($E \leq$ 1e−04) to 209 different *C. elegans* genes. These findings suggested that our dataset was 16.4% fragmented, or under-clustered; however, this should be considered a maximum value for reasons already presented and shown in Table 2. We also determined that the mean length of sequences

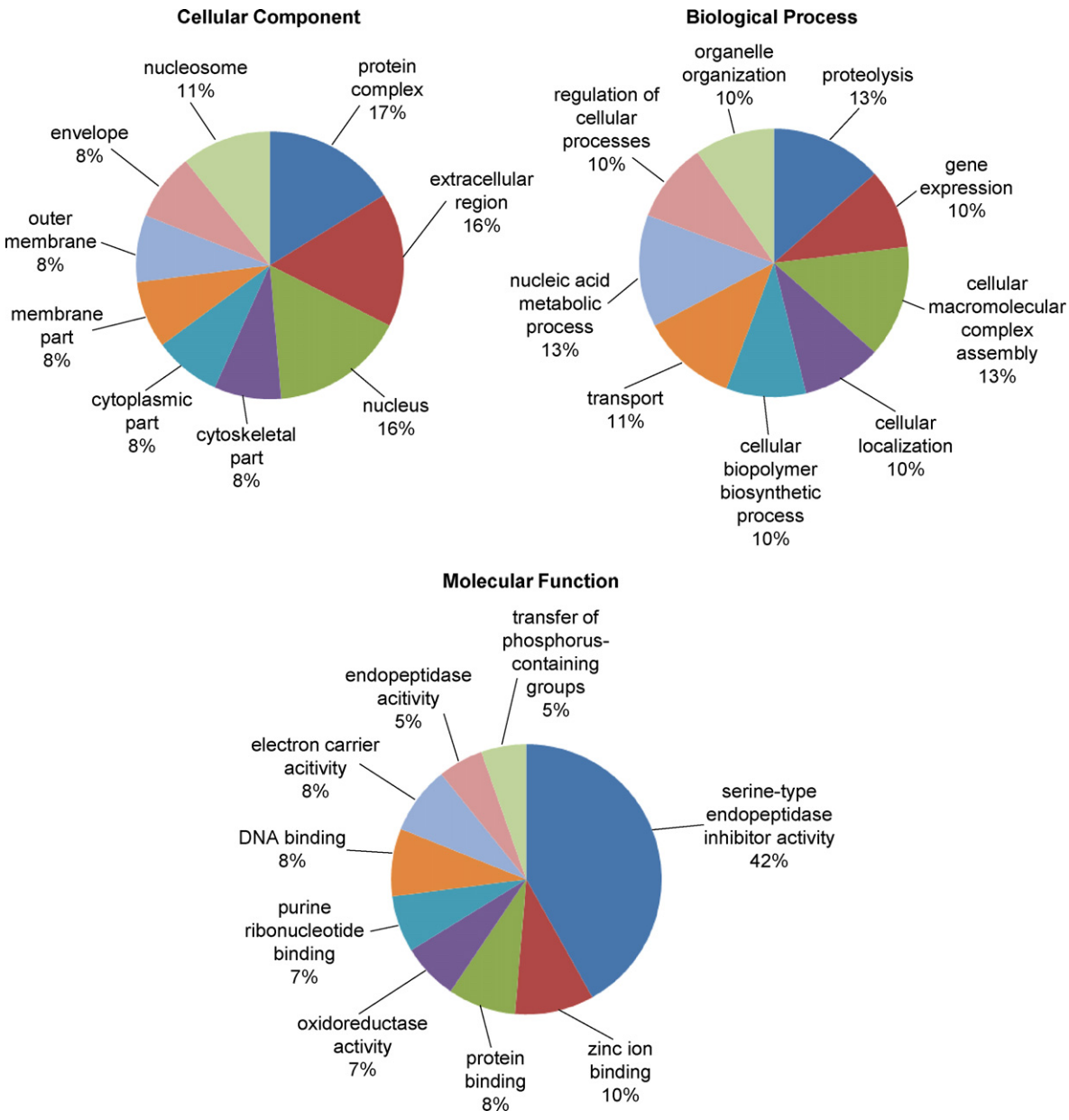

Fig. 2. Gene Ontology (GO) mappings for all unigenes as determined by Blast2GO. Pie charts correspond to the general categories of molecular function, biological process, and cellular component with sequence number cut-offs of 4, 5, and 3, respectively.

that showed similarity to a *C. elegans* homolog (554.5 nt) was significantly greater that the mean length of sequences which gave no hits (420.0 nt).

RNA-interference (RNAi) of some nematode genes has been shown to be effective in reducing plant-parasitic nematode infection and/or reproduction [15,23]. In some instances, RNAi-mediated silencing of plant-parasitic nematode genes that are homologous to *C. elegans* genes that show a lethal phenotype when silenced has been successful [15,19,42]. Using the *E*-value cut-offs of Alkharouf et al. [19], *R. reniformis* unigenes were divided into six groups (I–VI) based on the level of identity shared with their corresponding *C. elegans* homolog. Within each group, the number of *C. elegans* homologs having lethal and non-lethal RNAi phenotypes was determined (Fig. 4). Group I ($E \leq 1e{-}100$) contained only one sequence, FRC-7C_F08, which matched a *C. elegans* actin-related protein (Table 6). Group II ($1e{-}100 < E \leq 1e{-}80$) contained nine unigenes with six having *C. elegans* orthologs that show RNAi-lethality (Fig. 4 and Table 6). As *E*-values became increasingly less stringent across Groups III–VI, the number of unigenes within a group increased while the percentage of those unigenes within a group having RNAi-lethal *C. elegans* orthologs decreased (Fig. 4). Based upon this analysis, the 10 *R. reniformis* unigenes that showed the greatest potential for causing lethality upon silencing based on homology with *C. elegans* are presented in Table 6. The 10 unigenes represent proteins that are involved in conserved eukaryotic processes such as transcription, translation, glycolysis, and the citric acid cycle. FRC-4D_M13F_C07 is a RACK-1 homolog that has been shown to be required for cytokinesis in *C. elegans* [43].

4. Discussion

Plant-parasitic nematodes generally fall into one of four broad categories based on whether their feeding strategy is sedentary or migratory and whether feeding occurs inside or outside of the root. The reniform nematode, *R. reniformis*, is unusual in this regard as it is one of only a few sedentary semi-endoparasitic species. In

Table 4
Rotylenchulus reniformis consensus sequences with BLASTx matches only to parasitic genera in the Nematoda phylum.

Consensus sequence	Accession No.	BLASTx of consensus sequence	*E*-Value	SignalP[a]	TMHMM[b]
Contig103	EZ421805	gi|170587220|ref|XP_001898376.1|corneal wound healing-related protein [*Brugia malayi*]	1.7e−05	Negative	Negative
Contig116	EZ421812	gi|30315078|gb|AAP30768.1|putative gland protein G22C12 [*Heterodera glycines*]	6.6e−10	Positive	Negative
Contig46	EZ421897	gi|5457299|emb|CAA70477.2|SEC-2 protein [Globodera pallid]	8.1e−10	Negative	Negative
FRC-1A_M13F_E04	GT736555	gi|170578013|ref|XP_001894227.1|prolyl oligopeptidase family protein [*Brugia malayi*]	1.8e−32	Negative	Negative
FRC-2C_M13R_G08	GT736997	gi|170591200|ref|XP_001900358.1|transformer-2a3 [*Brugia malayi*]	1.4e−04	Positive	Negative
FRC-3A_H04	GT737081	gi|2463096|emb|CAA74206.1|serine proteinase [*Heterodera glycines*]	9.7e−12	Negative	Negative
FRC-5A_M13F_D07	GT737643	gi|37725601|gb|AAO25506.1|cellulase [*Heterodera glycines*]	5.9e−08	Negative	Negative

[a] Predicted peptide tested positive or negative for a signal peptide according to SignalP 3.0 [55].
[b] Predicted peptide tested positive or negative for the presence of transmembrane helices according to TMHMM 2.0 [56].

addition to being a possible evolutionary intermediate between sedentary endo- and ectoparasitism, *R. reniformis* is a notorious plant pathogen that is capable of infecting more than 300 plant species, including cotton, soybean, chickpea, and sweet potato [1]. As an alternative to natural host resistance, many research programs have focused on identifying and characterizing genes from many plant-parasitic nematode species that cause the parasite to die upon silencing by RNA-interference (RNAi) [23]. The molecular interaction between *R. reniformis* and its host has not been studied; consequently, no gene sequence data exists for this parasite. We present herein an initial survey of the *R. reniformis* transcriptome by identifying ESTs from parasitic sedentary females. This particular life-stage was chosen for EST analysis for the following reasons: (i) sedentary females represent the sole feeding life-stage, (ii) sedentary females are easily identified and collected compared to the non-feeding juvenile life-stages, (iii) *R. reniformis* genes involved in feeding site formation/maintainence should be expressed in sedentary females, and (iv) the potential success of future RNAi-based control strategies hinge upon the identification of those genes expressed by the nematode during feeding.

BLASTX was used to compare the 631 *R. reniformis* unigenes to the NCBI non-redundant database. We determined that 246 unigenes (39%) aligned with a known protein sequence at a maximum *E*-value threshold of 1e−04. Homology searches for other plant-parasitic nematodes have yielded similar results. For example, 46% of all *R. similis* ESTs yielded significant BLASTX matches at $E \leq$ 1e−05 [24]. Likewise, only 57% of sequences from the migratory endoparasite *D. africanus* showed similarity to known sequences [25]. Given the lack of sequences from other sedentary semi-endoparasites for comparison, it may not be surprising that more *R. reniformis* sequences appear to be unique relative to other plant-parasitic nematodes. The vast majority of unigenes that showed significant BLASTX matches identified with homologous sequences from the model bacteriovore *Caenorhabditis* and/or from the well-studied filarial parasite *B. malayi*. Two unigenes corresponded to known nematode-specific gene families, i.e., the fatty acid- and retinol-binding proteins (FAR) and the transthyretin-like proteins.

FAR proteins are particularly interesting because they have the potential to disrupt host defense signaling [38]. The top BLASTX match for Contig46 was the *G. pallida* FAR protein, SEC-2; however, subsequent BLASTX matches were almost exclusively to FAR proteins from filarial nematodes. This result may indicate that Contig46 represents a FAR protein involved in parasitism as opposed to a more general function like that hypothesized for some FAR proteins [24,44]. Six additional unigenes only matched proteins

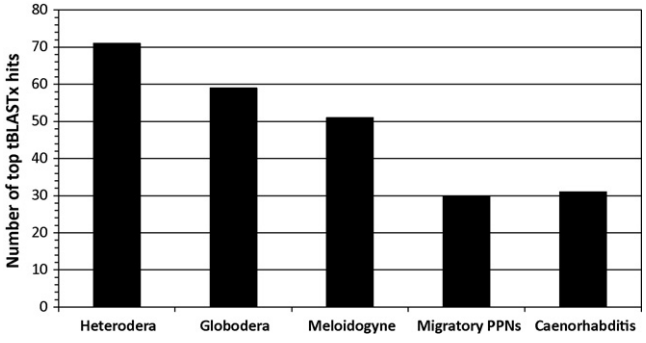

Fig. 3. Distribution of *Rotylenchulus reniformis* unigene top TBLASTX matches to the NCBI est_others database according to nematode genera. Migratory plant-parasitic nematodes (PPNs) refer to *Radopholus*, *Pratylenchus*, and *Bursaphelenchus* genera.

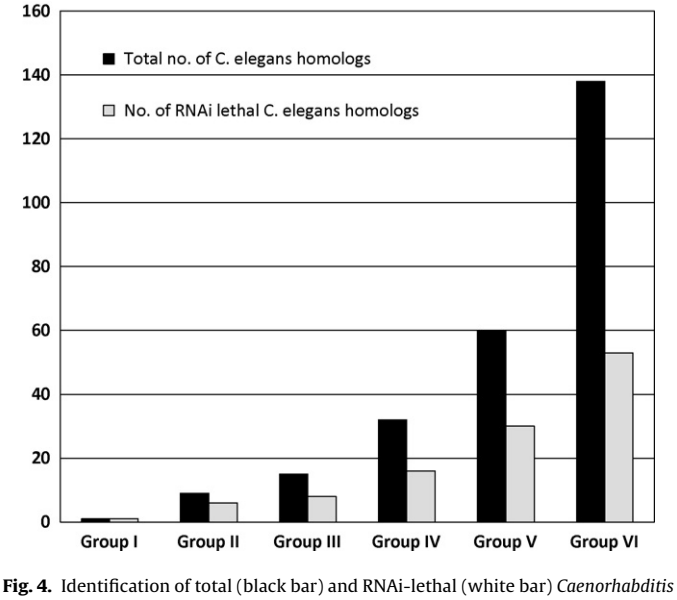

Fig. 4. Identification of total (black bar) and RNAi-lethal (white bar) *Caenorhabditis elegans* proteins homologous at different levels of significance to the *Rotylenchulus reniformis* unigene set. Group I ($E \leq$ 1e−100), Group II (1e−100 < $E \leq$ 1e−80), Group III (1e−80 < $E \leq$ 1e−60), Group IV (1e−60 < $E \leq$ 1e−40), Group V (1e−40 < $E \leq$ 1e−20), and Group VI ($E >$ 1e−20).

Table 5
Rotylenchulus reniformis consensus sequences that do not have BLASTn or BLASTx matches, lack predicted transmembrane domains, and test positive for a signal peptide.

Consensus sequence	GenBank accession number	Length (nt)	Reading frame/peptide length (aa)	tBLASTx of consensus sequence	BLASTp of predicted peptide
Contig12	EZ421815	485	−1/103	No hits	No hits
Contig67	EZ421911	464	+2/83	No hits	No hits
Contig154	EZ421826	1310	+2/296	rf44c02.y1 Meloidogyne hapla J2 pAMP1 v1 cDNA (2.6e−05)	No hits
Contig205	EZ421838	428	+1/81	No hits	No hits
Contig319	EZ421867	990	−3/264	No hits	No hits
Contig320	EZ421868	703	+2/178	gi|40305549|gb|CK349936.1| hggfha20B02 Gland Cell LD PCR cDNA Library *Heterodera glycines* cDNA (4.2e−05)	gb|AAO85454.1| putative gland protein G16B09 [*Heterodera glycines*] (7.0e−07)
Contig385	EZ421882	594	−3/149	No hits	No hits
Contig413	EZ421890	963	+1/270	gi|125428713|gb|EL440685.1| Jerusalem artichoke Helianthus tuberosus cDNA clone CHTM1715 (9.3e−04)	No hits
Contig436	EZ421893	631	+1/161	No hits	No hits
Contig472	EZ421899	766	−2/102	gi|29960829|gb|CB824593.1| rq28f06.y1 *Heterodera glycines* J3 *Heterodera glycines* cDNA (1.0e−10)	Cyt-b5, Cytochrome b5-like Heme/Steroid binding (4.0e−04)
FRC-01-test_M13F_E10	GT736489	277	−2/58	No hits	No hits
FRC-01-test_M13F_H03	GT736502	254	−3/>59[a]	No hits	No hits
FRC-1D_M13F_C11	GT736738	509	−2/80	No hits	No hits
FRC-1D_M13F_F12	GT736760	284	−2/40	No hits	No hits
FRC-2A_H03	GT736794	477	−2/73	No hits	No hits
FRC-2A_M13R_G10	GT736854	410	−2/77	No hits	No hits
FRC-2B_M13F_E10	GT736896	307	−2/62	No hits	No hits
FRC-2C_M13F_B02	GT736946	234	−1/>53[a]	No hits	No hits
FRC-2D_M13F_E03	GT737029	422	+3/88	No hits	No hits
FRC-3A_M13R_G04	GT737133	600	+2/72	gi|46422446|gb|CN443812.1| re69e08.y1 Meloidogyne incognita J3 J4 SL1 pGEM cDNA (1.4e−10)	No hits
FRC-4D_D09	GT737557	472	−2/84	No hits	No hits
FRC-4D_G02	GT737558	737	−2/175, +1/54[b]	No hits	No hits
FRC-5A_M13F_F09	GT737647	343	+1/78	No hits	sp|P85213.2|DEF1_GALME RecName: Full = Defensin; AltName: Full = Galiomicin (3.0e−05)
FRC-5C_M13F_F01	GT737815	439	+3/>137[a]	No hits	No hits
FRC-6B_M13F_F09	GT738016	497	+1/103	No hits	No hits
FRC-6B_M13F_F12	GT738018	702	−2/120	gi|240974642|gb|GR367707.1| Gpa_EST_upregulated_J16E6_A12_048 Globodera pallida J2 (8.3e−05)	No hits
FRC-6C_M13F_E09	GT738080	279	+3/65	gi|51237561|gb|CO897771.1| pe06h01.y1 *Radopholus similis* SL1 TOPO v1 cDNA (2.7e−04)	No hits
FRC-7C_M13F_G05	GT738369	400	+3/94, −1/68[b]	No hits	No hits
FRC-7D_M13R_D08	GT738447	529	−1/112	No hits	No hits

[a] Open reading frame containing the signal peptide lacked a STOP codon.
[b] Signal peptide was detected (probability >0.900) for proteins derived from two open reading frames.

Table 6
Ten most homologous *Caenorhabditis elegans* genes that show RNAi-lethal phenotypes.

Consensus sequence	GenBank accession no.	*C. elegans* homolog	Homolog description	*E*-Value
FRC-7C_F08	GT738331	WBGene00000199	Actin-related protein Arp2/3 complex, subunit Arp 3	1e−135
Contig119	EZ421814	WBGene00001167	Translation elongation factor 2	9e−94
FRC-01-test_D07	GT736461	WBGene00009126	Pyruvate kinase	4e−93
FRC-6C_M13F_F05	GT738084	WBGene00001209	Eukaryotic translation initiation factor 3, subunit 10	1e−88
FRC-5C_H04	GT737786	WBGene00003162	NAD-dependent malate dehydrogenase	6e−87
FRC-6A_G12	GT737931	WBGene00021845	DNA-directed RNA polymerase, subunit E′	8e−85
FRC-3D_A03	GT737270	WBGene00000151	Apurinic/apyrimidinic endonuclease/3′-repair diesterase	7e−82
FRC-4D_M13F_C07	GT737570	WBGene00010556	RACK1 (mammalian receptor of activated C kinase homolog)	2e−76
FRC-2C_M13F_D02	GT736955	WBGene00008546	Glucosamine-fructose 6-phosphate aminotransferase	8e−69
FRC-4A_E02	GT737339	WBGene00020837	Amino acid transporter	9e−68

from other parasitic nematode genera. This list included Contig116 which showed homology to the *H. glycines* cDNA clone G22C12, a parasitism gene expressed in the dorsal esophageal gland cell of parasitic second-stage and third-stage juveniles [11].

Some unigenes showed significant BLASTX matches to only non-nematode sequences. Of these, Contig218 was particularly interesting because it showed a high degree of similarity to a homolog of NMT-1 (No Message in Thiamin-1) from the fungus *G. zeae*. Thiamin (vitamin B1) is an essential nutrient for all animals, including nematodes [45]; however, the ability to synthesize thiamin *de novo* is restricted to bacteria, fungi, and plants [46,47]. Evidence suggests that in yeast NMT-1, also known as THI5, is responsible for the synthesis of the thiamin precursor HMP-PP (hydroxymethylpyrimidine diphosphate) [48]. A functional NMT-1 homolog in *R. reniformis* would be the first report of its kind within the kingdom Animalia. Also, TBLASTX analysis of Contig218 revealed that homologous sequences exist for the potato cyst nematode *G. rostochiensis*; however, no other nematode-related hits were identified. A possible explanation for the existence of a NMT-1 homolog in *R. reniformis* would be that the nematode acquired this gene through the process of horizontal gene transfer (HGT). HGT has been used to explain the existence of cell wall degrading enzymes in PPN genomes [49,50] as well as the presence of a seemingly intact vitamin B6 biosynthesis pathway in *H. glycines* [51].

The use of RNAi to silence PPN genes required for feeding site formation or for conserved cellular processes has shown tremendous promise towards developing resistance in crop hosts where natural resistance is non-existent, insufficient, or difficult to incorporate into elite germplasm [23]. PPN genes targeted for silencing have included those expressed only in the esophageal gland cells, i.e., parasitism genes [15–17,52,53], and those believed to be generally necessary for survival [15,19–22,54]. It has become standard practice to search for suitable RNAi targets for a given species based on the identification of *C. elegans* homologs that show severe RNAi-mediated phenotypes. For example, 1508 *H. glycines* genes were identified that were conserved in *C. elegans* and showed a lethal phenotype when silenced or mutated [19]. In our analysis of 631 *R. reniformis* unigenes, 255 (40%) showed significant identity ($E \leq 1e{-}05$) to genes from *C. elegans*. Of these 255 conserved *C. elegans* genes, 114 had been shown to produce severe RNAi-mediated phenotypes; however, only 10 showed a level of identity where $E < 1.0e{-}60$. Silencing of a *R. reniformis* ribosomal RNA may be possible as was demonstrated in *H. glycines* [15,19]. In many instances, a crop species is a suitable host to more than one kind of nematode; cotton is an excellent host for both reniform nematode and root-knot nematode. It will be interesting to see if a gene can be identified that is conserved between *R. reniformis* and *M. incognita* at a level sufficient to enable its silencing in both species simultaneously from a single RNAi construct.

The unigene sequences presented in this report represent a small portion of the entire *R. reniformis* transcriptome. Despite this fact, many interesting sequences were discovered, including a short list of candidate parasitism genes that can be tested for gland cell-specific expression in the nematode. Future EST studies will implement next-generation-sequencing technology which will result in a more complete understanding of gene expression in the reniform nematode. This sequence data helps fill an information void in plant-parasitic nematology and provides a starting point for studying *R. reniformis* from a functional genomic perspective.

Acknowledgements

The authors would like to thank Drs. John Brooks, Mike McLaughlin (USDA/ARS) and Vincent Klink (Mississippi State University) for their critical reviews of the manuscript and suggestions. Approved for publication as Journal Article No. J-11671 of the Mississippi Agricultural and Forestry Experiment Station, Mississippi State University.

References

[1] Robinson AF. Reniform nematodes: *Rotylenchulus* species. In: Starr JL, et al., editors. Plant resistance to parasitic nematodes. CAB International; 2002. p. 153–74.
[2] Gaur HS, Perry RN. The biology and control of the plant parasitic nematode *Rotylenchulus reniformis*. Agric Zool Rev 1991;4:177–212.
[3] Blasingame D. 2005 cotton disease loss estimate committee report. 2006; 2006 Beltwide Cotton Conferences, San Antonio, TX, USA, January 3–6.
[4] Koenning SR, Kirkpatrick TL, Starr JL, Wrather JA, Walker NR, Mueller JD. Plant-parasitic nematodes attacking cotton in the United States. Plant Disease 2004;88:100–13.
[5] Starr JL, Koenning SR, Kirkpatrick TL, Robinson AF, Roberts PA, Nichols RL. The future of nematode management in cotton. J Nematol 2007;39:283–94.
[6] Robbins RT, Rakes L, Jackson LE, Dombek DG. Reniform nematode resistance in selected soybean cultivars. Suppl J Nematol 2001;31:667–77.
[7] Davis EL, Hussey RS, Baum TJ. Getting to the roots of parasitism by nematodes. Trends Parasitol 2004;20:134–41.
[8] Davis EL, Hussey RS, Mitchum MG, Baum TJ. Parasitism proteins in nematode–plant interactions. Curr Opin Plant Biol 2008;11:360–6.
[9] Vanholme B, de Meutter J, Tytgat T, Van Montagu M, Coomans A, Gheysen G. Secretions of plant-parasitic nematodes: a molecular update. Gene 2004;332:13–27.
[10] Lambert KN, Allen KD, Sussex IM. Cloning and characterization of an esophageal-gland-specific chorismate mutase from the phytoparasitic nematode *Meloidogyne javanica*. Mol Plant-Microbe Interact 1999;12:328–36.
[11] Gao B, Allen R, Maier T, Davis EL, Baum TJ, Hussey RS. The parasitome of the phytonematode *Heterodera glycines*. Mol Plant-Microbe Interact 2003;16:720–6.
[12] Lu SW, Chen S, Wang J, et al. Structural and functional diversity of CLAVATA3/ESR (CLE)-like genes from the potato cyst nematode *Globodera rostochiensis*. Mol Plant-Microbe Interact 2009;22:1128–42.
[13] Elling AA, Mitreva M, Gai X, et al. Sequence mining and transcript profiling to explore cyst nematode parasitism. BMC Genomics 2009;10:58.
[14] Fire A, Xu SQ, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 1998;391:806–11.
[15] Klink VP, Kim K-H, Martins V, et al. A correlation between host-mediated expression of parasite genes as tandem inverted repeats and abrogation of development of female *Heterodera glycines* cyst formation during infection of *Glycine max*. Planta 2009;230:53–71.
[16] Chen Q, Rehman S, Smant G, Jones JT. Functional analysis of pathogenicity proteins of the potato cyst nematode *Globodera rostochiensis* using RNAi. Mol Plant-Microbe Interact 2005;18:621–5.
[17] Huang G, Allen R, Davis EL, Baum TJ, Hussey RS. Engineering broad root-knot resistance in transgenic plants by RNAi silencing of a conserved and essential root-knot nematode parasitism gene. Proc Natl Acad Sci USA 2006;103:14302–6.
[18] Sindhu AS, Maier TR, Mitchum MG, Hussey RS, Davis EL, Baum TJ. Effective and specific in planta RNAi in cyst nematodes: expression interference of four parasitism genes reduces parasitic success. J Exp Bot 2009;60:315–24.
[19] Alkharouf NW, Klink VP, Matthews BF. Identification of *Heterodera glycines* (soybean cyst nematode [SCN]) cDNA sequences with high identity to those of *Caenorhabditis elegans* having lethal mutant or RNAi phenotypes. Exp Parasitol 2007;115:247–58.
[20] Bakhetia M, Charlton W, Atkinson HJ, McPherson MJ. RNA interference of dual oxidase in the plant nematode *Meloidogyne incognita*. Mol Plant-Microbe Interact 2005;18:1099–106.
[21] Lilley CJ, Goodchild SA, Atkinson HJ, Urwin PE. Cloning and characterization of a *Heterodera glycines* aminopeptidase cDNA. Int J Parasitol 2005;35:1577–85.
[22] Shingles J, Lilley CJ, Atkinson HJ, Urwin PE. *Meloidogyne incognita*: molecular and biochemical characterization of a cathepsin L cysteine proteinase and the effect on parasitism following RNAi. Exp Parasitol 2007;115:114–20.
[23] Rosso MN, Jones JT, Abad P. RNAi and functional genomics in plant parasitic nematodes. Annu Rev Phytopathol 2009;47:207–32.
[24] Jacob J, Mitreva M, Vanholme B, Gheysen G. Exploring the transcriptome of the burrowing nematode *Radopholus similis*. Mol Genet Genomics 2008;280:1–17.
[25] Haegeman A, Jacob J, Vanholme B, Kyndt T, Mitreva M, Gheysen G. Expressed sequence tags of the peanut pod nematode *Ditylenchus africanus*: the first transcriptome analysis of an Anguinid nematode. Mol Biochem Parasitol 2009;167:32–40.
[26] Kikuchi T, Aikawa T, Kosaka H, Pritchard L, Ogura N, Jones JT. Expressed sequence tag (EST) analysis of the pine wood nematode *Bursaphelenchus xylophilus* and *B. mucronatus*. Mol Biochem Parasitol 2007;155:9–17.
[27] Kang JS, Lee H, Moon IS, et al. Construction and characterization of subtractive stage-specific expressed sequence tag (EST) libraries of the pinewood nematode *Bursaphelenchus xylophilus*. Genomics 2009;94:70–7.
[28] Mitreva M, Elling AA, Dante M, et al. A survey of SL1-spliced transcripts from the root-lesion nematode *Pratylenchus penetrans*. Mol Genet Genomics 2004;272:138–48.

[29] Barker KR. Nematode extraction and bioassays. In: Barker KR, et al., editors. An advanced treatise on *Meloidogyne*, vol II, methodology. Raleigh, NC: North Carolina State University Graphics; 1985. p. 19–38.
[30] Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 1998;8:175–85.
[31] Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res 1998;8:186–94.
[32] Chou H-H, Holmes MH. DNA sequence quality trimming and vector removal. Bioinformatics 2001;17:1093–104.
[33] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215:403–10.
[34] Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 2005;21:3674–6.
[35] Reboul J, Vaglio P, Tzellas N, et al. Open-reading-frame sequence tags (OSTs) support the existence of at least 17,300 genes in *C. elegans*. Nat Genet 2001;27:227–8.
[36] Abad P, Gouzy J, Aury J-M, et al. Genome sequence of the metazoan plant-parasitic nematode Meloidogyne incognita. Nat Biotechnol 2008;26:909–15.
[37] Popeijus H, Blok VC, Cardle L, et al. Analysis of genes expressed in second stage juveniles of the potato cyst nematodes *Globodera rostochiensis* and *G. pallida* using the expressed sequence tag approach. Nematology 2000;2:567–74.
[38] Prior A, Jones JT, Blok VC, et al. A surface-associated retinol- and fatty acid-binding protein (Gp-FAR-1) from the potato cyst nematode *Globodera pallida*: lipid binding activities, structural analysis and expression pattern. Biochem J 2001;356:387–94.
[39] Jurgenson CT, Begley TP, Ealick SE. The structural and biochemical foundations of thiamin biosynthesis. Annu Rev Biochem 2009;78:569–603.
[40] Mitreva M, McCarter JP, Martin J, et al. Comparative genomics of gene expression in the parasitic and free-living nematodes *Strongyloides stercoralis* and *Caenorhabditis elegans*. Genome Res 2004;14:209–20.
[41] Roze E, Hanse B, Mitreva M, Vanholme B, Bakker J, Smant G. Mining the secretome of the root-knot nematode *Meloidogyne chitwoodi* for candidate parasitism genes. Mol Plant Pathol 2008;9:1–10.
[42] Yadav BC, Veluthambi K, Subramaniam K. Host-generated double stranded RNA induces RNAi in plant-parasitic nematodes and protects the host from infection. Mol Biochem Parasitol 2006;148:219–22.
[43] Ai E, Poole DS, Skop AR. RACK-1 directs dynactin-dependent RAB-11 endosomal recycling during mitosis in *Caenorhabditis elegans*. Mol Biol Cell 2009;20:1629–38.
[44] Garofalo A, Kennedy MW, Bradley JE. The FAR proteins of parasitic nematodes: their possible involvement in the pathogenesis of infection and the use of *Caenorhabditis elegans* as a model system to evaluate their function. Med Microbiol Immunol 2003;192:47–52.
[45] de Jong L, Meng Y, Dent J, Hekimi S. Thiamine pyrophosphate biosynthesis and transport in the nematode *Caenorhabditis elegans*. Genetics 2004;168:845–54.
[46] Begley TP, Downs DM, Ealick SE, et al. Thiamin biosynthesis in prokaryotes. Arch Microbiol 1999;171:293–300.
[47] Morett E, Korbel JO, Rajan E, et al. Systematic discovery of analogous enzymes in thiamin biosynthesis. Nat Biotechnol 2003;21:790–5.
[48] Wightman R, Meacock PA. The *THI5* gene family of *Saccharomyces cerevisiae*: distribution of homologues among the hemiascomycetes and functional redundancy in the aerobic biosynthesis of thiamin from pyridoxine. Microbiology 2003;149:1447–60.
[49] Dieterich C, Sommer RJ. How to become a parasite – lessons from the genomes of nematodes. Trends Genet 2009;25:203–9.
[50] Mitreva M, Smant G, Helder J. Role of horizontal gene transfer in the evolution of plant parasitism among nematodes. Methods Mol Biol 2009;532:517–35.
[51] Craig JP, Bekal S, Hudson M, Domier L, Niblack T, Lambert KN. Analysis of a horizontally transferred pathway involved in vitamin B6 biosynthesis from the soybean cyst nematode *Heterodera glycines*. Mol Biol Evol 2008;25:2085–98.
[52] Bakhetia M, Urwin PE, Atkinson HJ. qPCR analysis and RNAi define pharyngeal gland cell-expressed genes of *Heterodera glycines* required for initial interactions with the host. Mol Plant-Microbe Interact 2007;20:306–12.
[53] Rosso MN, Dubrana MP, Cimbolini N, Jaubert S, Abad P. Application of RNA interference to root-knot nematode genes encoding esophageal gland proteins. Mol Plant-Microbe Interact 2005;18:615–20.
[54] Urwin PE, Lilley CJ, Atkinson HJ. Ingestion of double-stranded RNA by preparasitic juvenile cyst nematodes leads to RNA interference. Mol Plant-Microbe Interact 2002;15:747–52.
[55] Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP3.0. J Mol Biol 2004;340:783–95.
[56] Krogh A, von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. J Mol Biol 2001;305:567–80.